我国科研人员对 arXiv 认知和使用的现状调查与分析

张智雄；张闪闪；顾立平；李麟

中国科学院文献情报中心 北京 100190

摘要：【目的】了解中国用户对 arXiv 预印本平台的认知、使用情况以及对 arXiv 发展的意见建议。**【方法】**基于国内 9 家高校及科研机构的教师、研究人员和研究生群体的问卷调查，作者获得 510 份调查数据，并运用 SPSS 等分析工具对调查数据进行分析。**【结果】**结果表明，我国科研人员对 arXiv 的认知程度还不是很高，但熟悉 arXiv 的国内科研人员已认识到在 arXiv 上发布论文是占据成果首发权和征求同行意见的重要手段。**【局限】**在调查取样上仅限于中国 arXiv 服务工作组的成员单位，非成员单位的 arXiv 用户没有纳入。**【结论】**为了让 arXiv 能够更多地惠及我国科研，中国 arXiv 服务工作组需要采取更为积极的行动，加快推动 arXiv 在国内更广泛深入的应用。

关键词： ArXiv 科研人员 用户调查 认知度 使用意愿

分类号：G250

1. 研究背景与目的

作为一个预印本论文存缴、检索、发布和交流共享服务的开放存储平台，arXiv 受到信息技术专家、图书馆员、科研管理者和科研人员的高度关注，为了协助建设者和管理者更好地促进 arXiv 在中国的建设和发展，本次 arXiv 用户调查针对科研教育机构中的研究人员以及研究生群体对 arXiv 认知度、使用意愿、投稿意愿、使用情况和使用中的问题。

量子物理学家 Paul Henry Ginsparg 于 1991 年创建、在 2001 年转由 Cornell University 进行维护和管理的 arXiv 主要提供物理学、数学、计算机科学、计量生物学、计量金融学、统计学领域的预印本论文的存缴、检索、发布和交流共享服务，是重要的开放存储系统[1]。科研人员将论文按照一定的排版格式整理后，通过 Web 和 E-mail 等方式，按学科类别上传至相应的 arXiv 数据库中，提供开放共享、避免重复工作、增加信息交流效率，在国际领域的科学家团体中具有重要的影响。国外的著名案例是俄罗斯数学家格里戈里·佩雷尔曼（Grisha Perelman）在本世纪初对庞加莱猜想的证明成果首先在 arXiv 上发布 [2]。我国的著名案例是在 2012 年大亚湾中微子实验国际合作组在大亚湾中微子实验发现新的中微子振荡，也首先在 arXiv 上发布[3]。有人统计截止到 2012 年 10 月，它已经存储了 79 万余篇研究论文[4]；在 2014 年 3 月底，预印本电子文献数量已达 927,838 篇[5]；在将近一年半的时间里，共新增 13 万篇左右的论文，平均每月存储 8000 余篇，相比 2011-2012 年增长 76578 篇[6]文献来看，即每月新增文献约 6380 篇左右；截止到 2013 年 12 月底，每月的下载量也几乎达到了 550 万篇左右[7]。

随着科研用户对 arXiv 的认识，学者们对 arXiv 的研究也在逐渐增多，这也为今后 arXiv 网站的资源组织、功能完善等提供了有用的借鉴意义[8]。目前已有学者对 arXiv 展开了各种研究，比如 2007 年，Davis, Philip M.[9]分析了四种数学期刊中的 2765 篇文献，发现 arXiv 增加了论文的引用率，但一定程度上减少了从出版商网站下载文献的数量；Moed, Henk F.[10]分析了开放获取对期刊引用影响力的影响效果，发现 arXiv 主要是使论文可以被更早地获取，而不是可免费获取；2009 年，Mine, Shinji[11]分析了 arXiv 在学术交流中的角色和地位，研究结果表明用户向 arXiv 提交论文的数量不断上升等结果。在 2010 年，我国的孙玉伟[12]分析了 arXiv 的路径模型，并从研究角度、研究内容、研究方法、研究的深度和广度四个方面和 Hurd 模型、Sφndergaard 模型、SCLC 模型等科学交流模型进行了对比分析；2011 年王欣[13]等分析了国内物理学期刊 arXiv 自存档论文，通过定量的数据统计，发现通过 arXiv，论文的被引频率、下载次数都有所增加，建议国内物理学期刊出版者加强网站建设，充分利用

arXiv 平台提升期刊引用水平；孙斌[14]以 arXiv 为例剖析了预印本的运行机制，通过统计数据分析了其在学术信息交流中的作用，文章最后指出预印本的使用与 Ingelfinger 编辑政策、研究峰值之间的关系；王佳佳[15]对 ArXiv 和国内预印本系统的投入与利用情况、资源数量和作者的分布进行研究，从侧面反映出目前国内的预印本系统的发展现状及存在的问题，并提出相应的建议对策；2013 年张智雄[16]等探讨了 arXiv 在科研成果开放发表中的特殊地位，分析了 arXiv 在中国科技界的主要应用情况，并推出了 2012-2013 年中国 arXiv 工作组的具体工作计划等。

为了充分理解中国用户的使用情况与改进建议，延续自 1998 年开始的中国镜像站.arXiv.org[17]服务，中国 arXiv 服务工作组进行在 2013 年 9 月期间实施了中国用户对 arXiv 知识库的调查，调查对象主要是工作组的 9 家成员单位[一]（不包含中国科学院文献情报中心）的 500 多名科研用户。

2. 研究方法与过程

本次调查对象的科研用户主要包括高校、科研机构中的教师、研究人员以及研究生群体，针对中国用户对 arXiv 知识库的使用情况、使用习惯以及对 arXiv 发展的意见建议等进行问卷调查，以期为 arXiv 知识库在中国的发展提供借鉴意义。

2.1. 问卷内容设计

本次调查组共设计 18 个问题（回答这些问题大概需要 10 分钟，具体问卷详见文后）。本次调查主要包含两个方面，一是国内科研用户对 arXiv 知识库的认知情况，主要目的在与了解不同单位、不同年龄、不同学科领域、不同身份乃至不同学历的用户对 arXiv 知识库的总体认知；二是调查国内科研用户对 arXiv 知识库的使用情况，包括使用方式、投稿目的、投稿后的影响、向 arXiv 投稿的影响因素、浏览 arXiv 方式、检索过程中遇到的问题等。问卷的 18 个问题包括单选题、跳转题、多序列题以及开放性问题。此外，还有用户的身份信息，包括所属学科领域、性别、年龄、职称、学历。

2.2. 问卷调查方式

考虑到 arXiv 目前在国内的发展，而中国 arXiv 服务工作组成员单位的科研用户基本可以代表国内教育和科研机构的主要成员类型，而且成员单位自身也拥有大量的具有一定实力的科研用户。因此，本次调查面向中国 arXiv 服务工作组的 9 家成员单位，将问卷发放给各成员单位的相关负责人，由各成员单位负责人进行选择、确定本机构内部的科研用户进行问卷调研，同时每个科研机构在选择调查对象时，也要结合用户的学历、职称、学科领域等，以保证调查结果的可靠性。

2.3. 问卷的发放与回收

本次调研采取纸质版和电子版相结合的形式进行。纸质问卷由各成员单位负责人进行当面发放与回收，所有问卷都经过统一编码后，寄送中国科学院文献情报中心。电子问卷是通过电子邮件进行点对点的收集，再由成员单位负责人进行统一汇总编码后发送至中国科学院文献情报中心。两类问卷汇总后进行统一处理和分析。此次调研中，向服务工作组的 9 所单位进行调研，要求每单位提供 50 份问卷，在后期调查阶段，扩大样本规模为 540 份，最终回收问卷 510 份，回收率为 94.4%，其中有效问卷 509 份，有效率为 99.8%。其中纸质问卷 146 份，占有效问卷总数的 28.7%，电子问卷 363 份，占总数的 71.3%。

2.4. 样本数据的有效性分析

本次研究首先利用 SPSS 中的可靠性分析来分析问卷的信度，进而利用 KMO 和 Bartlett 的检验来分析其效度，如图 1 所示。由于问卷包含 3 道跳转题，因此分析时进行了区分。对

一 本文的 9 家调查单位分别为：中国科学院理论物理研究所、清华大学图书馆、北京大学图书馆、上海交通大学图书馆、复旦大学图书馆、南京大学图书馆、中国科学院高能物理研究所、吉林大学图书馆、中国科学院数学与系统科学研究院。

本次调查问卷的信度进行分析。

可靠性统计量

Cronbach's Alpha	基于标准化项的 Cronbachs Alpha	项数
.942	.937	14

可靠性统计量

Cronbach's Alpha	基于标准化项的 Cronbachs Alpha	项数
.953	.956	17

图 1 信度分析（排除及包含跳转题的两种情况）

从图 1 可以看出，无论是排除跳转题还是包括跳转题，Cronbach's Alpha 的值都大于 0.9，这说明问卷具有很高的内在信度，表明问卷的内在一致性非常好，很适合进行后续的分析。

3. 调查结果与讨论

3.1. 被调查对象分析

从调查问卷的回收情况来看，此次被调查对象主是要年龄分布在 45 岁以下的中青年科研人员，占总数的 93.3%，其中又以 20-25 岁年龄段的科研人员所占比例最高，超过了三分之一，达 36%。被调查对象中，还有部分年龄在 45-55 岁和 55 岁以上的科研人员，应当说他们是相关领域的资深科研人员。从受教育程度上来看，被调查对象中以有博士学位或正在读博士学科的科研人员为多，占总量的 65.8%，硕士/硕士研究生次之，为 22.4%，大学生的比例最少。从专业职称上看，在读研究生、副研究员和助理研究员是本次被调查对象的主体。其中在读研究生占总数的 57.3%，副研究员为 16.3%，助理研究员为 13.9%。从学科分布来看，本次被调查对象的学科领域主要为高能物理、凝聚态物理、天体物理、数学、计算机、计量生物学、计量经济学、统计学学科，其中数学领域和高能物理领域所占比例较大，分别为 25.9%和 20.4%，其后依次是凝聚态物理、计算机、天体物理、统计学、计量经济学、生物学。

3.2. 认知和使用意愿分析

（1）人们知晓 arXiv 的途径主要来自同事同学的推举

在本次问卷调查的 509 个样本中，知道 arXiv 的样本量为 342 个，比重占 67.2%，虽然超过了一半，但是从 arXiv 作为高能物理、凝聚态物理、天体物理、数学、计算机、计量生物学、计量经济学、统计学领域内一种很重要的学术资源来看，这一比例并不高。调查发现，用户对 arXiv 的知晓途径主要是通过同事、同学的推荐。被调查者中，通过本机构图书馆的宣传、培训而得知 arXiv 的所占比例较小，仅仅为 14.9%。如图 2 所示：

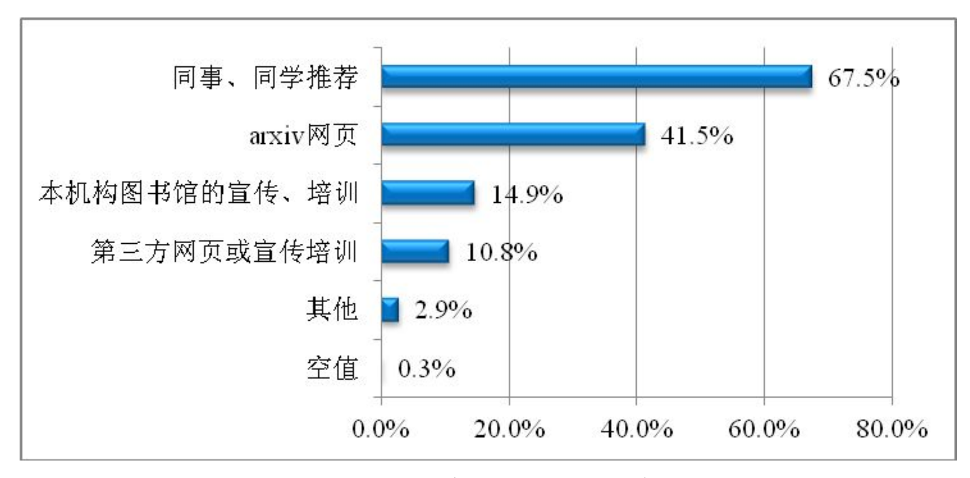


图 2　用户对 arXiv 的知晓途径

（2）用户使用 arXiv 的方式以下载文献为主，浏览为辅

知道 arXiv 的 342 个被调查者中，有 82.6%的被调查者使用过 arXiv，未使用过的达 16.3%，这说明了知道 arXiv 的被调查者大部分都使用过。在使用过 arXiv 的被调查者中，科研人员使用 arXiv 主要是为了下载文献。如图 3 所示。

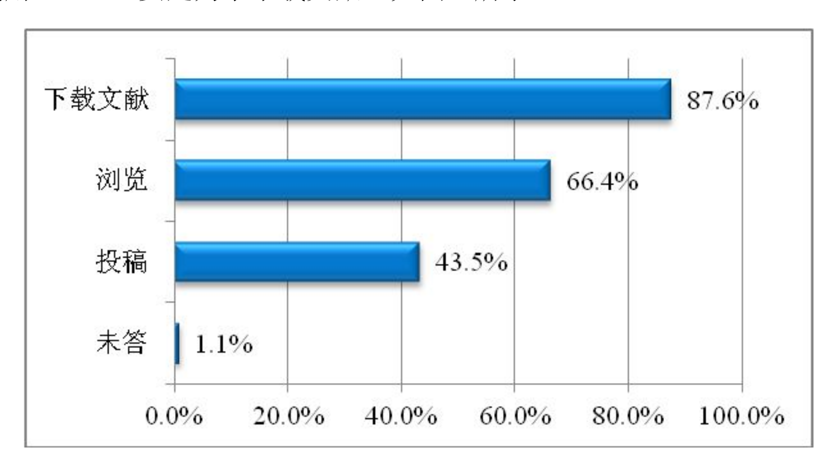


图 3　用户使用 arXiv 的方式

（3）用户知晓但是却未使用 arXiv 的原因，部分来自 arXiv 不能简明易懂，部分来自图书馆缺乏推广培训

本次调查，知晓但是却未使用 arXiv 文档库的用户共有 56 名，其中 arXiv 网站缺乏清楚简捷易懂的使用说明是用户未使用 arXiv 的主要原因。此外，用户所在的图书馆缺乏相应的培训课程和咨询帮助、arXiv 内容与我的学科领域不相关两个因素对影响用户使用 arXiv 的比例相差不大，分别为 26.8%和 23.2%，这也从另一方面进一步验证了宣传和推广 arXiv 的重要性。

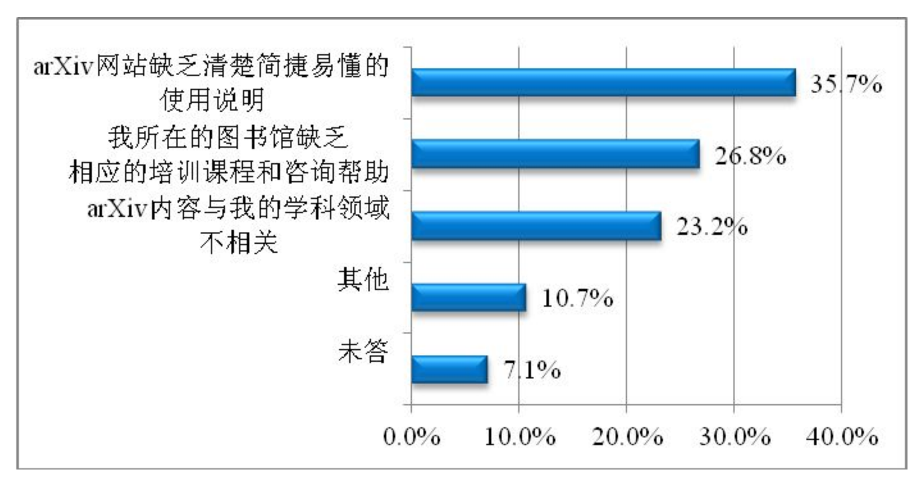


图 4　用户知晓但没有使用 arXiv 的原因

3.3. 投稿意愿分析

（1）多数用户偶尔进行投稿，并不会以 arXiv 为每稿必投的对象

在本次调查中，共有 123 个被调查者向 arXiv 提交过论文（投稿）。在投稿频率上，偶尔投的最多，占 39.0%，每稿必投的其次，比例为 35.8%，经常投的占 21.1%。

用户在 arXiv 投稿的数量如图 5 所示：以 1-5 篇的居多，占 61.0%。

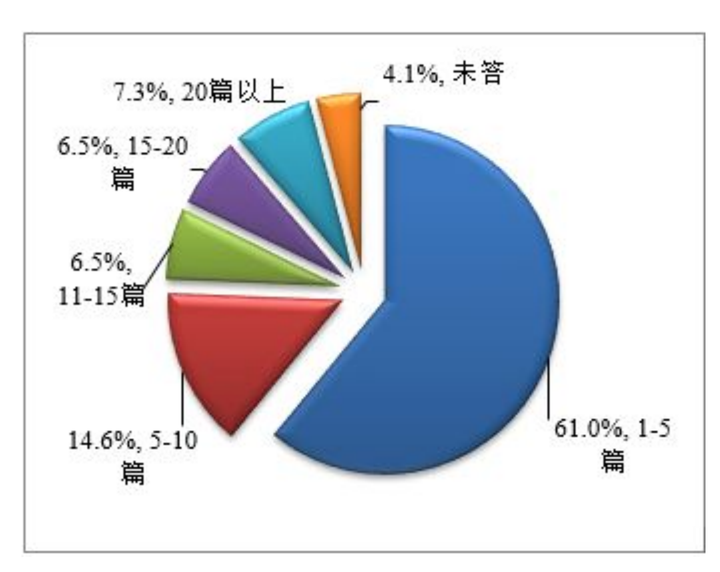


图 5　用户在 arXiv 投稿的数量

（2）向 arXiv 投稿的目的主要是占据成果的首发权和征求更多的同行意见

在 123 个向 arXiv 提交过论文的被调查者，提交的论文大多是未发表的论文。基于投稿目的的调查数据，可以非常明显地看到，这些被调查者之所以向 arXiv 投稿，主是占据成果的首发权和征求更多的同行意见。

这些 123 个被调查者中，在 arXiv 的著作权要求的理解上，大多数被调查者（73.1%）认为自己已经了解了 arXiv 对论文著作权的要求，但还有一部分被调查者（22.0%）认为不了解 arXiv 对论文的著作权要求。

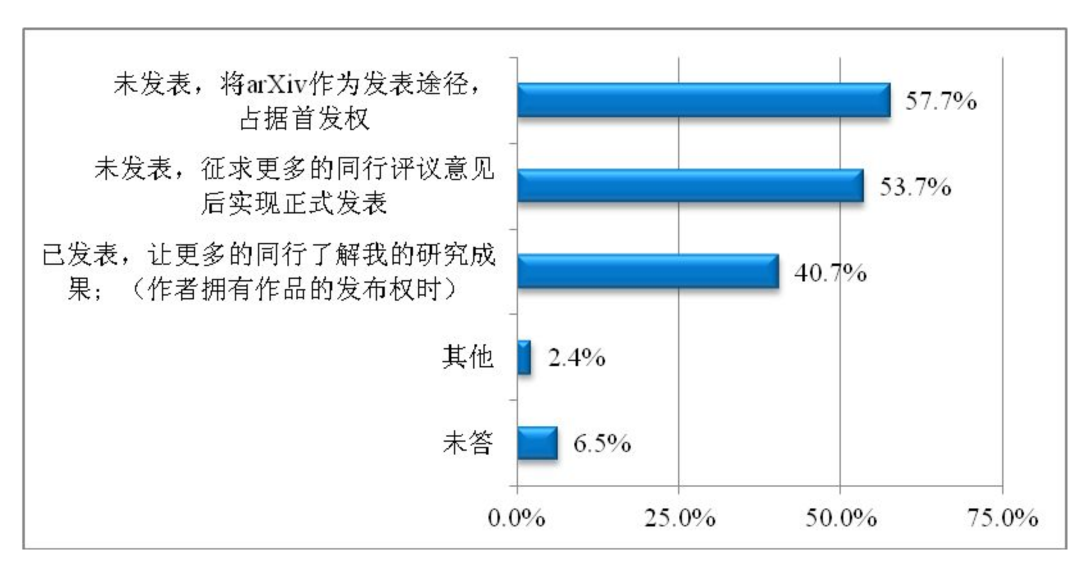


图 6　用户向 arXiv 投稿的目的

（3）通过向 arXiv 投稿，科研人员能够结识更多的同行，并提高论文被引率

从分析结果来看，在本次调查的 123 个通过投稿方式使用 arXiv 的样本中，用户认为，通过向 arXiv 投稿，使他们结识了更多的研究同行的比例最高，为 53.7%，其次为论文的被引率提高，比例为 44.7%，而对于论文是否更容易在期刊上发表的影响并不是太大，比例仅为 17.1%。

（4）影响用户在 arXiv 投稿的五项因素

据调查显示，促使科研人员愿意将其科研成果上传至 arXiv 的前五名影响因素是发布的速度、得到更多同行的意见、arXiv 的整体访问量和投稿量、同行意见反馈的速度、投稿步骤操作简便。如图 7 所示。

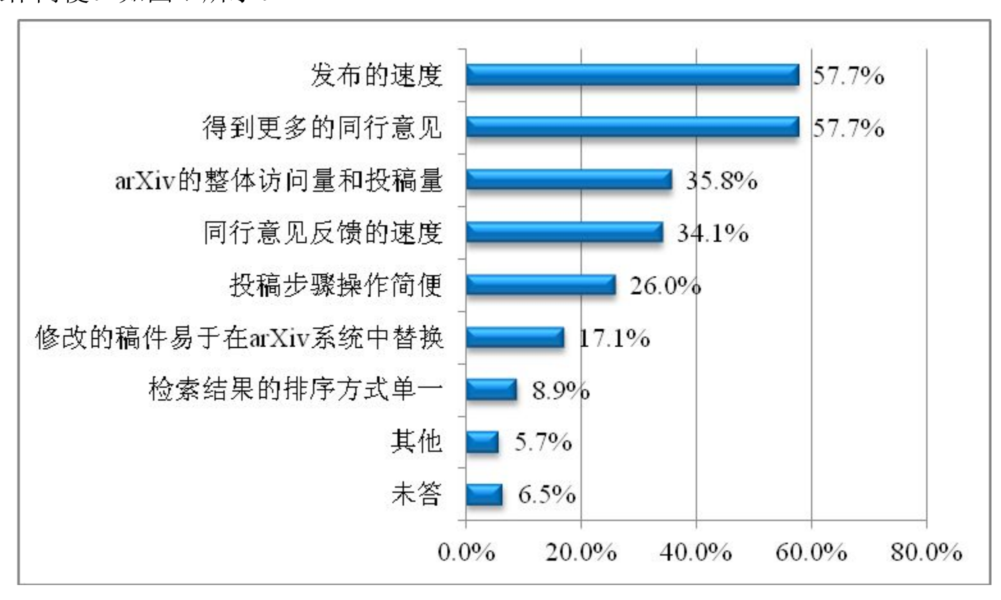


图 7　影响用户在 arXiv 投稿的因素

3.4. 使用情况及使用问题分析

（1）有 91.1%的科研用户愿意将自己的科研成果上传至 arXiv

调查结果显示，总体上来讲，约有 91.1%的科研用户愿意将自己的科研成果上传至 arXiv，有 7.4%的用户明确表示不愿意上传，但也有 1.5%的用户没有明确表达的意愿。

（2）浏览 arXiv 的方式及时间

从调查结果来看，有 92.6%的用户通过网页方式浏览 arXiv，而利用邮件（Email alerts）

提醒功能和 RSS 订阅新文章功能的方式浏览 arXiv 的比例较低，分别为 8.5%和 6.4%。

从用户浏览 arXiv 的时间来看，每周浏览一次的用户比较多，比例为 31.8%，每天浏览 arXiv 的用户所占的比例为 25.4%，每月浏览的比例为 18.0%，每两周浏览的比例为 10.2%，还有 12.4%的用户是在其他时间浏览 arXiv 的。

（3）不使用 arXiv 的原因主要是检索困难

在本次调查的 283 个使用 arXiv 的样本中，有 138 个人没有 arXiv 使用上的问题，占 48.8%，但也有 109 个人认为检索中存在比较大的问题，占 38.5%，至于 arXiv 的浏览上，仅有 7.1%的用户认为有问题，此外，还有 6.0%的用户认为在使用过程中存在一些其他方面的问题，比如系统出现问题后，更新的速度比较慢。

（4）检索困难的主要问题是缺乏高级检索和检索不准确

结果表明，在本次调查的 283 个使用 arXiv 的样本中，检索入口太少，缺少高级检索功能、检索结果不准确、检索结果增加文章的质量统计指标这三个问题是用户在检索过程中经常遇到的问题，他们所占的比例分别为 34.6%，32.9%，24.0%，而检索结果的数量、反馈速度上，用户并没有太多的问题，比例仅为 10.6%和 9.5%如图 8 所示：

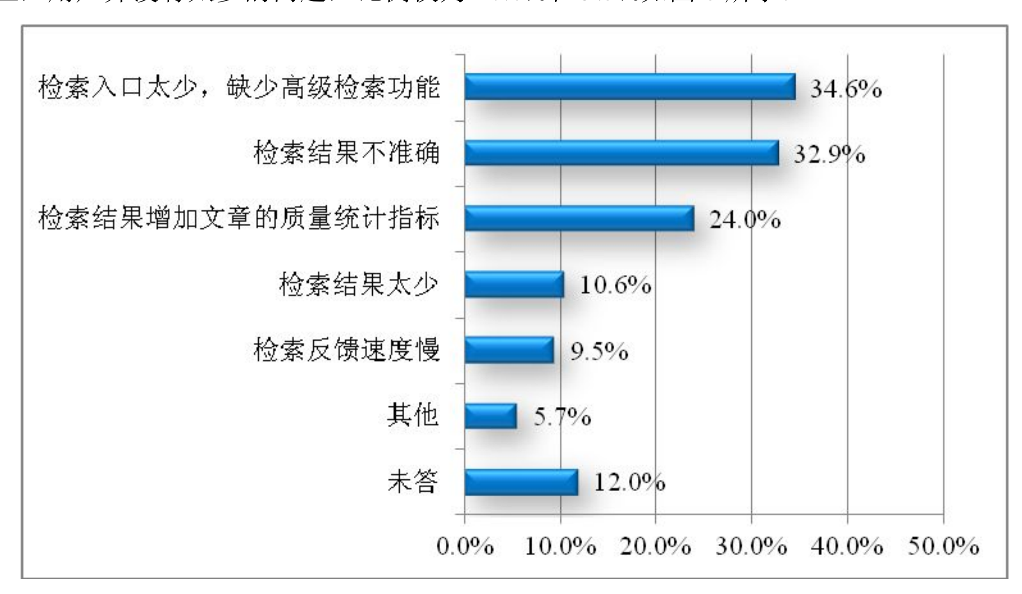


图 8 用户在 arXiv 进行检索时遇到的问题

（5）arXiv 服务还有其他小问题

如图 9 所示，在本次调查的 283 个使用 arXiv 的样本中，还提供了其它的问题反馈。

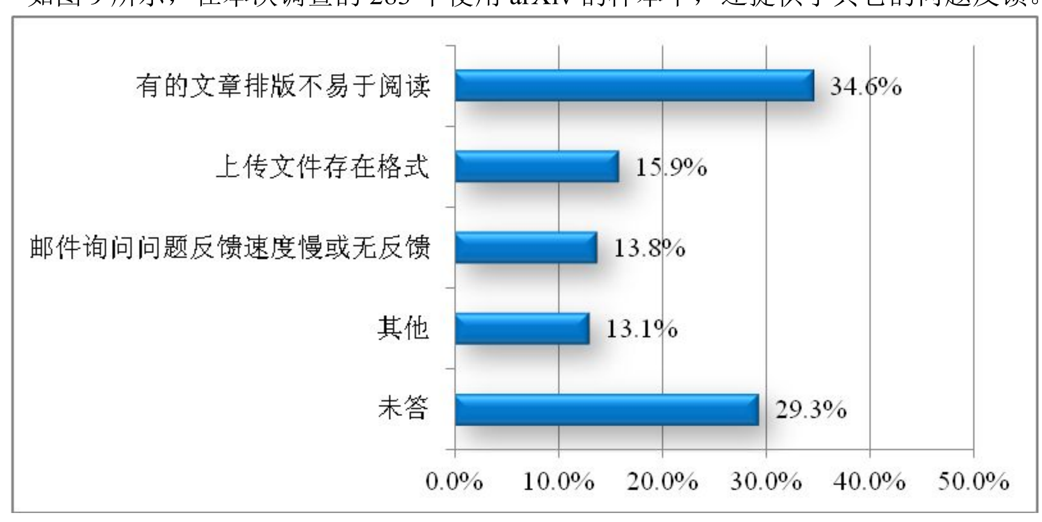


图 9 用户使用 arXiv 过程中的其他问题

4. 研究结论与建议

本次对我国科研用户的 arXiv 知识库的使用情况调研是国内首次大规模的关于 arXiv 的调查，以 arXiv 服务工作组 9 家成员单位的科研用户为调查对象，不仅有科研机构，也有高等教育用户的参与；不仅有学生，也有研究员/教授、副研究员/副教授、助理研究员/讲师，年龄阶段也从 20 岁到 55 岁以上不等。基本上能够反应绝大多数的国内用户对 arXiv 的认知和使用情况。

4.1. 主要发现

通过上述调查工作，笔者发现：

（1）我国科研人员对 arXiv 的认知程度还不是很高。科研人员主要通过同事、同学们的推荐而知道 arXiv，图书馆在 arXiv 的宣传培训上存在一定的缺失。

（2）我国科研人员已经认识到 arXiv 在学术交流中发挥着重要作用。熟悉 arXiv 的科研人员已认识到在 arXiv 上发布论文是占据成果首发权和征求同行意见的重要手段；通过向 arXiv 投稿，科研人员能够结识更多的同行，并提高论文被引率。

（3）arXiv 在系统平台方面的问题影响了用户的使用。平台的检索功能不完善、检索结果不准确是用户们普遍反应的问题。

4.2. 研究限制

进行研究推论时，需要注意到本次调查针对中国 arXiv 服务工作组成员单位内的用户，这些用户也应当是我国 arXiv 应用的主要用户，但本调查的确没有将中国 arXiv 服务工作组成员单位以外的 arXiv 用户包括在内。

另一方面，本文的分析结果仅仅是调查问卷的统计结果，由于条件所限，作者没能与被调查者进行更为深入地交流互动，也没有进行更为深入的案例调研分析。

4.3. 研究建议与未来工作

通过对本次 arXiv 使用情况的调查，可以认为，尽管 arXiv 的浏览量、下载量都在不断上升，但是我国科研用户真正了解 arXiv 的还在少数，大多用户对 arXiv 还缺乏相应的认知。一些用户还紧紧处于只浏览，不投稿的态度，这可能是由于用户对 arXiv 不太了解，也可能是由于其对著作权的疑虑和担心。虽然 arXiv 的建设是一个长期的工程，但是基于本次的调研，笔者认为，为了让 arXiv 能够更多地惠及我国科研，中国 arXiv 服务工作组需要采取更为积极的行动，加快推动 arXiv 在国内更广泛深入的应用。

（1）加大 arXiv 的广泛宣传

在现有调查群体中，只有 67.2%的用户知道 arXiv，而通过本机构图书馆的宣传以及第三方网页或宣传两种方式来了解 arXiv 的用户仅为 10%左右，因此，大力宣传 arXiv 是现阶段工作人员的重要任务，arXiv 工作组成员应该制定良好的宣传计划，在本校或科研机构内部开展 arXiv 的宣传讲座，或在网页设置链接、公告栏介绍等，鼓励更多的用户认识 arXiv，提升认知度。除此之外，工作组成员还应该多多鼓励 arXiv 的现有用户，让现有用户通过正式和非正式渠道进行宣传，以用户带动用户的方式，挖掘 arXiv 的潜在价值，引导科研用户正确认识 arXiv，积极发挥 arXiv 的学术交流作用。

（2）加强 arXiv 的培训引导

缺乏有效的培训，是高校和科研机构用户没有使用 arXiv 的主要原因之一。加强有效的培训引导是提高 arXiv 使用率的条件之一，科研机构要加大对 arXiv 培训的资金投入，在机构内部设立相关的培训课程，形成培训大纲和培训材料，广泛推行 arXiv 的使用培训，引导用户正确使用。在用户认识 arXiv 的基础上，具体讲解 arXiv 如何使用、投稿规范、检索策略、对著作权的规定等，这就需要在今后很长一段时间内，将工作重心转向 arXiv 使用的培训和相关问题的解决上，积极动员中国的广大科研用户积极参与内容存缴，帮助用户更好地了解 arXiv 网站，也可以邀请康纳尔大学图书馆等在此方面经验丰富的机构的相关人员针对

物理学、计算机、统计学等领域的科研用户展开深入的培训。

（3）完善 arXiv 的平台建设

在本次的调查分析结果中，arXiv 网站缺乏清楚简捷易懂的使用说明是影响用户使用 arXiv 的最主要原因，占到了 35.7%，而投稿步骤的简捷性、文章的排版问题也是影响用户在 arXiv 上投稿的主要因素，arXiv 平台的规范与完善，直接影响用户的使用。要尽快通过定期的使用情况调查，了解物理学、数学、统计学等学科用户的使用反馈意见，根据用户反馈的特点完善 arXiv 平台建设，加强其投稿功能的设计，改善检索功能，并优化相关服务，提升用户在 arXiv 投稿的效率。针对于浏览问题，可以设置清楚的标引，建立使用说明文档，引导用户正确使用；对于投稿，可以改进投稿步骤，设置简便的投稿流程，明确规定投稿的版本、版权、利用和再利用的政策，促进更多用户投稿。

（4）改善 arXiv 的用户体验

在信息环境时代，任何一个数据库或者平台库要想得到用户的重用，一方面需要有大量的资源，另一方面也需要有良好的用户服务体验，arXiv 定位于提供物理学、化学、数学等领域的预印本文献的存缴、检索、发布，存储特定领域的知识资产和成果，然而科研人员在 arXiv 进行检索、获取信息时，由于不同科研人员所处的环境，以及 arXiv 的宣传培训，用户的使用体验不尽相同。因此，要加大对 arXiv 内容质量的建设，为用户提供优质、良好的信息资源；借鉴 MY Yahoo![18]、谷歌的 AdWards 平台[19]以及百分点信息科技有限公司的百分点分析引擎[20]的成功个性化推荐经验，既可以通过网站设立个性化信息服务的入口，也可以通过加强 RSS 和 E-alerts 功能，为用户订阅属于自己的专项数据库；要加强 arXiv 平台与搜索引擎、商业数据库等的无缝链接，改善用户体验。

（5）重视 arXiv 的有效激励

ArXiv 的建设与各方的积极性密不可分，相关领域的机构单位要鼓励科研用户将其科研成果上传至 arXiv，为科研用户创造论文开放共享的环境，鼓励其更加积极地参与到 arXiv 的学术交流中去；要支持科研机构的科研成果和论文的开放共享，要支持科研机构及其科研资助机构将其工作重心转移到提高机构、科研成果的社会影响力上；相关机构也可以完善科研绩效策略，比如可以将论文的引用量和科研用户的绩效挂钩，一方面可以促进科研用户积极参与到 arXiv 中，另一方面也可以加快 arXiv 的建设步伐；期刊也应该积极接收经过同行评议的，已上传至 arXiv 文档库的论文的正式发表，保障知识的广泛传播与共享。

（6）规范 arXiv 的流程机制

部分文章的排版不容易阅读是目前用户在使用 arXiv 中的最大问题，因此，在 arXiv 的建设、推广和使用过程中要制定详细的投稿流程、管理机制和为用户的服务机制等。比如在宣传和培训上，形成统一、标准、清晰的 arXiv 宣传培训材料，形成相关的统一课件，提供给工作组成员单位以及其他需要进行 arXiv 培训宣传的单位进行使用；对平台建设上，增加使用说明，设置清晰标引，规范论文格式，给用户利用上的最大便利；对 arXiv 的管理机制、服务机制，也要设立明确的平台管理标准，提升平台的利用率。

致谢：中国 arXiv 用户调查是在“中国 arXiv 服务工作组”（http://chinair.csdl.ac.cn/node/.）的规划、组织和指导下实施，感谢工作组成员为本次的调查提供了可贵的数据。本文部分内容曾在 2013 年中国开放获取推介周上进行汇报，感谢与会专家的交流指教。

参考文献：

[1] Ginsparg P. As we may read[J]. The Journal of Neuroscience, 2006,26(38):9606-9608.
[2] Wikipedia. Grigori perelman.[EB/OL].[2014-03-28]. http://en.wikipedia.org/wiki/Grigori_Perelman.
[3] Observation of Electron Anti-neutrino Disappearance at Daya Bay.

[EB/OL].[2012-03-08].http://www.ihep.cas.cn/xwdt/gnxw/2012/201203/t20120308_3450968.html.
[4] ArXiv [EB/OL].[2012-11-19].http://arxiv.org/pdf/math/0303109.
[5] ArXiv [OL].[2014-03-31].http://arxiv.org/.
[6] David Ruddy. Sustaining Open Access Repositories.[EB/OL].[2012-10-22]. http://ir.las.ac.cn/handle/12502/5495.
[7] arXiv monthly download rates.[OL].[2014-03-31].http://arxiv.org/stats/monthly_downloads.
[8] Zhang Zhixiong. arXiv.org China Service Group. 2012 China Open Access Week [EB/OL]. [2012-11-23]. http://ir.las.ac.cn/handle/12502/5496.
[9] Davis, Philip M.; Fromerth, Michael J.. Does the arXiv lead to higher citations and reduced publisher downloads for mathematics articles?[J]. SCIENTOMETRICS.2007,71(02):203-215.
[10] Moed, Henk F.. The effect of "Open access" on citation impact: An analysis of ArXiv's condensed matter section.[J].JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY. 2007,58(13):2047-2054.
[11] Mine, Shinji.The Roles and Place of arXiv in Scholarly Communication[J].LIBRARY AND INFORMATION SCIENCE. 2009(61):25-58.
[12] 孙玉伟. 数字环境下科学交流模型的分析与评述[J].大学图书馆学报,2010,01.41-45.（Sun Yuwei. Analysis on Scientific Communication Models in Digital Environments[J]. Journal of Academic Libraries, 2010,01.41-45.）
[13] 王欣,董洪光.国内物理学期刊 arXiv 自存档论文的引用优势研究——以 Frontiers of Physics 期刊为例[J].图书情报工作.2011,22:144-148.（Wang Xin, Dong Hongguang. Citation Advantage of arXiv Self-archiving Articles Published in Domestic Physics Journals—Taking Frontiers of Physics as an Example[J]. Library and Information Service, 2011,22:144-148.）
[14] 孙斌.预印本在网络环境下的复兴与发展[J].图书馆工作与研究.2011,06:70-73.（Sun Bin. On the Revival and Development of the Preprints in Network Environment[J].Library Work and Study. 2011,22:144-148.）
[15] 王佳佳.Arxiv 与国内预印本系统的比较研究[J].科技情报开发与经济.2011,27:104-107.（Wang Jiajia. A Comparative Study on Arxiv and Domestic Preprint System[J]. Sci-Tech Information Development & Economy. 2011,27:104-107.）
[16] 张智雄;顾立平;张晓林;李麟. 组建中国 arxiv 服务工作组促进我国有效参与 OA 活动[J].图书情报工作,2013,57(1):55-59.（Zhang Zhixiong, Ku liping, Zhang Xiaolin, Li Lin. Establishing Arxiv China service group to promote open access movement in China[J]. Library and Information Service,2013,57(1):55-59.）
[17] cn.arXiv.org[OL].[2012-03-28].http://cn.arxiv.org/.
[18] my.yahoo![OL].[2014-04-05].https://my.yahoo.com/.
[19] AdWords[OL].[2014-04-05]. https://accounts.google.com/ServiceLogin?service=adwords&continue=https://adwords.google.com/um/gaiaauth?apt%3DNone%26ltmpl%3Djfk&hl=zh_CN&ltmpl=jfk&passive=86400&skipvpage=true&sacu=1&sarp=1&sourceid=AWO&subid=CN-ET-ADS.
[20] Baifendian Analytics Engine [OL].[2014-04-05]. http://www.baifendian.com/list.php?catid=64.

作者贡献声明：

张智雄，顾立平：提出调查思路，组织设计调查方案；

李麟，张智雄：调查问卷的设计与修改；

李麟，顾立平：组织实施调查；

张闪闪，顾立平：问卷数据分析与论文初稿；

张智雄，顾立平：负责论文最终版本的修订。

※ 本文排名不分先后，按照拼音字母笔画列出署名顺序：ZZX, ZSS,KLP,LL。

（通讯作者：张智雄 E-mail: zhangzhixiong@mail.las.ac.cn）

Survey and Analysis on Cognition and Using of ArXiv for China Mainland Researchers

Zhang Zhixiong；Zhang Shanshan；Ku Liping； Li Lin

(National Science Library, Chinese Academy of Sciences, Beijing 100190, China)

Abstract:[Objective] Understanding the cognition, conditions, habits on arXiv of China Mainland Researchers, and their opinions and suggestions for the long-term development of arXiv. **[Methods]** This paper takes the researchers of educational and research institutions as survey object, obtains the survey datum from more than 500 teachers, researchers and graduate students of 9 institutions. And, this paper analysis the datum on the base of the reliability and validity. **[Results]** 540 questionnaires were given out and 510 effective recalled, and the rate was 94.4%.according to this, the paper acquired the cognition, conditions and habits on arXiv of China Mainland Researchers, and their opinions and suggestions for the long-term development of arXiv in line with the survey. Also, based on the data, obtaining the cognition, submission, usage and retrieval differences between research institutions and universities communities. **[Limitations]** The objective users was just belong to the participating member, the users who outside the participating member didn't participate in this survey. The findings are just based on the survey data, they may not be applied to a wider range of user groups. **[Conclusions]** According to the survey, this paper brings out six suggestions on promoting arXiv development.

Keywords: ArXiv Researchers User survey Cognition Usage intention

arXiv 存档库用户调查问卷

您好！

衷心感谢您在繁忙的工作学习之余填写本调查问卷。arxiv.org 是世界上著名的预印本文献存档库，创建于 1991 年 8 月，是一个自动化程度很高的、活跃的研究论文的电子预印本存档和发布平台。主要覆盖的学科领域有：物理、数学、计算机、计算生物学、计量经济学和统计学。

为使中国用户更好地利用这一资源，经与 arxiv 协商，在中国建立了 arxiv 镜像站，并于 2012 年 10 月，由中国科学院国家科学图书馆和清华大学图书馆联合国内十余家大学和研究机构图书馆牵头成立了中国 arxiv 工作组，面向国内用户宣传和培训这一资源。

本调查旨在了解中国用户对 arxiv 存档库的使用情况、使用习惯以及对 arxiv 发展的意见建议，并向 arxiv 反馈调查结果。您填写的调查内容对我们非常重要，填写问卷大约需要 5-10 分钟，感谢您的支持与配合！

本调查完全采取匿名方式，我们保证所有调查信息和数据仅作为调查结果分析使用，不会用于商业等其他用途。

中国 arxiv 工作组

2013 年 9 月 16 日

请在下列各题要选择的选项前☐内划勾，或按各题括号中的提示填写。

一、arxiv 的认知和使用行为调查

1. 您是否知道 arxiv？

☐知道；　☐不知道（请跳至第 18 题）

2. 您从哪些途径知道的 arxiv？（可多选）

☐arxiv 网页

☐第三方网页或宣传培训

☐本机构图书馆的宣传、培训

☐同事、同学、老师推荐

☐其它（请说明）____________________________

3. 您是否使用过 arxiv？（可多选）

☐使用过（请跳至第 5 题）　☐没用过

4. 您知道 arxiv 但没有使用的原因是？（可多选）

☐arxiv 网站缺少清楚简捷易懂的使用说明

☐我所在的图书馆缺乏相应的培训课程和咨询帮助

☐arxiv 内容与我的学科领域不相关

☐其它（请说明）________________________

5. 您使用 arxiv 的方式？（可多选）

☐投稿（请回答第 6-12 题）　☐下载文献（请跳至第 12 题）

☐浏览（请跳至第 12 题）

6. 您在 arxiv 投稿的频率？（单选）

☐每稿必投　☐经常投　☐偶尔投

7. 您在 arxiv 投稿的论文数量？（单选）

☐1-5 篇　☐5-10 篇　☐11-15 篇　☐15-20 篇　☐20 篇以上

8. 您在 arxiv 投稿的目的：（可多选，并对您的选项按重要性 1、2、3……排序，1 为最重要）

☐未发表，征求更多的同行评议意见后实现正式发表；

☐未发表，将 arxiv 作为发表途径，占据首发权；

☐已发表，让更多的同行了解我的研究成果；（作者拥有作品的发布权时）

☐其它（请说明）_____________________

9. 您在 arxiv 投稿后对研究学习带来了什么影响？（可多选）

☐论文的被引率提高了

☐论文更容易在期刊上正式发表了

☐期刊不同意我正式发表该篇论文

☐认识了更多的研究同行

☐其它（请说明）____________________

10. 您投稿时是否了解 arxiv 对论文的著作权要求（注：arxiv 要求提交者授予 arxiv 论文的发布权）

☐了解 ☐不了解

11. 影响您在 arxiv 投稿的因素？（可多选，并对您的选项按重要性 1、2、3……排序，1 为最重要）

☐得到更多的同行意见；

☐发布的速度；

☐同行意见反馈的速度；

☐投稿步骤操作简便；

☐修改的稿件易于在 arxiv 系统中替换；

☐arxiv 的整体访问量和投稿量；

☐检索结果的排序方式单一；

☐其它（请说明）____________________；

12. 您今后是否打算投稿（之前未投过稿）或继续向 arxiv 投稿（投过稿)？（单选）

☐是 ☐否

13. 您浏览 arxiv 的方式（可多选）

☐网页浏览；☐订阅 RSS；☐利用邮件提醒（email alerts）功能提醒阅读新文章

14. 您多长时间浏览 arxiv 一次？（单选）

☐每天；☐每周；☐每两周；☐每月；☐其它（请说明）__________

15. 使用 arxiv 的过程中，您遇到的问题是？（可多选）

☐检索问题；☐浏览问题；☐其它问题（请说明）________；☐没有问题；

16. 您检索 arxiv 遇到的问题是？（可多选）

☐检索反馈速度慢；

☐检索结果不准确；

☐检索结果太少；

☐检索入口太少，缺少高级检索功能；

☐检索结果增加文章的质量统计指标；

☐其它（请说明）_______________；

17. arxiv 服务的其它问题？（可多选）

☐邮件询问问题反馈速度慢或无反馈；
☐有的文章排版不易于阅读；
☐上传文件存在格式；
☐其它（请说明）________________；

18. 您对 arxiv 的意见与建议？（您的意见与建议对我们向 arxiv 提出功能和服务改进非常重要，请您给予宝贵意见。）

19. 为便于调查统计和进一步沟通，请您提供以下信息，所有信息仅用于问卷统计分析。

（1）研究领域所属学科（可多选）

☐物理（请选择：☐高能物理；☐凝聚态物理；☐天体物理；☐数学； ☐计算机； ☐计量生物学； ☐计量经济学； ☐统计学； ☐其它（请说明）___________________

（2）性别：☐男；☐女

（3）年龄：☐20-25 岁；☐25-30 岁；☐30-35 岁；☐35-45 岁；☐45-55 岁；☐55 岁以上

（4）职称：☐研究员/教授；☐副研究员/副教授；☐助理研究员/讲师；☐实习研究员/助教；☐学生

（5）学历：☐博士或博士研究生；☐硕士或硕士研究生；☐大学生

（6）email:____________________________________（如您希望了解调查结果或愿意就相关问题进一步交流，请留下 email）

感谢您的参与和支持，祝您工作学习愉快！